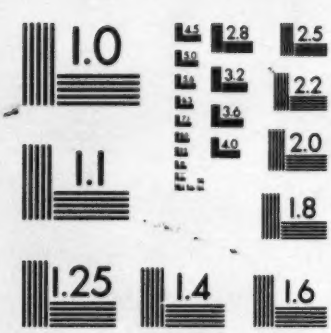

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

DOCUMENT RESUME

ED 215 657 HE 015 092

AUTHOR Lyons, Paul R.
TITLE Basing Performance Assessment on Behaviorally Anchored Rating Scales in Collegiate Organizations.
PUB DATE Apr 82
NOTE 16p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Administrator Evaluation; *Behavior Rating Scales; *Department Heads; *Evaluation Criteria; Higher Education; Performance; Research Methodology; Teacher Attitudes

ABSTRACT

The use of behaviorally anchored rating scales (BARS) as the basis of an assessment system that was designed to improve academic department chairpersons in a college of arts and sciences is described. Twenty-eight faculty members, two from each department, were asked to identify evaluative dimensions for assessing chairperson performance and to provide critical/behavioral incidents that would demonstrate poor, adequate, and good performance on each of the 11 dimensions. After review by a panel, 236 incidents were identified and were rewritten into an expectations format. Forty-two faculty were asked to make two judgments for the list of items. The first judgment required categorization of the 245 items into 11 performance dimensions. The second judgment required placement of each item on a seven-point scale based on the level of performance indicated. Seven scales that were generated were distributed to all faculty of 14 academic departments to generate primary information for a chairperson performance assessment. System application and examination of behaviorally anchored rating scale results, interviews with faculty and key administrative staff, and self-reports of chairpersons will be components of the chairperson performance assessment. Perspectives on administrative performance evaluation and features of BARS are also considered. It is suggested that the generation of BARS itself helps to clarify goals, and that its participative and collaborative characteristics help to ensure that values of the collegial body are being considered. A bibliography is appended. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED215657

BASING PERFORMANCE ASSESSMENT ON BEHAVIORALLY ANCHORED RATING SCALES IN COLLEGIATE ORGANIZATIONS

by

PAUL R. LYONS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Paul R. Lyons

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

APRIL, 1982

Graduate School

Frostburg State College

Frostburg, Maryland 21532

HE 015 072

BASING PERFORMANCE ASSESSMENT ON BEHAVIORALLY ANCHORED RATING SCALES IN COLLEGIATE ORGANIZATIONS

Introduction

The assessment of managerial performance in colleges and universities has not enjoyed wide success and little research has been reported regarding assessment of performance of academic administrators, in particular. Assessment as used in this paper includes activities carried out to enable the academic manager to improve his or her performance, as well as activities undertaken to determine quality of performance from the judgmental perspective of a senior manager or administrator.

The purpose of this paper is to describe how an organization has used behaviorally anchored rating scales (BARS) as the basis of an assessment system intended to provide information for self-improvement for academic department chairpersons in a college of arts and sciences.

Perspective

Cohen and March (1974) refer to the college or university organization as an organized anarchy which exhibits the following general properties: (1) problematic goals (ones that are vague or in dispute); (2) unclear technology (its own processes are not well understood); and, (3) fluid participation, where organization participants are free to vary in the amount of effort and time they devote to the organization, over time.

Within such a setting the assessment of performance is difficult because most models of assessment and evaluation are premised on theories of management and administration which assume the presence of well-defined goals as well as substantial participant involvement in the activities of the organization

(Cohen and March, p. 4). The methods of assessment used in the present study take special cognizance of the characteristics of the college organization with attempts to validate performance by way of multiple methods of assessment.

Using multiple methods to assess performance has been labeled triangulation (Green and Stone, 1977). Triangulation allows one to appraise the same variables from several aspects which may permit confirmation, substantiation, and verification of observations. Thus, multiple methods may compensate for inadequacies or weaknesses of a single method of assessment. In the present investigation the administrators of a college of arts and sciences used the methods of structured interviews with faculty, interviews with key administrative staff, self-reports and self-ratings of performance by the subjects and, finally, the empirically-derived, behaviorally anchored rating scales for use with all college of arts and sciences faculty. A thorough assessment effort could, as a positive aspect, lend credence to the overall act of assessment in the eyes of the subjects, academic department chairpersons.

The academic department chairperson occupies a pivotal position in the development and implementation of academic programs because he/she interacts with many institutional offices and members on a day-to-day basis regarding the delivery of educational services and because the proper and smooth operation of the academic program is dependent upon the performance of the department head or chairperson. The role and functions of chairpersons is examined extensively in the literature of higher education (Brown, 1977).

Evaluation of the administrative performance of the chairperson is largely a matter of concern internal to the university. Evaluation of chairpersons will most likely result in the opportunity for improvement in performance

through assessments of strengths and needs, and through an awareness of perceptions of persons with whom the chairperson works.

Systematic evaluation of administrative performance in universities has not enjoyed wide success. Farmer (1979) states that evaluation of administrative functions is a highly politicized process that contains little objectivity. Dressel (1976) warns that evaluation of administrative functions is most difficult in higher education organizations because few people agree on what criteria define success in administration. Booth (1978, p. 80) reports that many case studies have shown the capacity of chairpersons to make improvements in the operation of academic departments and he goes on to say that more systematic attention to the evaluation of chairpersons would probably produce a good many administrative improvements.

The role and functions of the academic department chairperson were carefully examined in a monograph by Waltzer (1975). Through detailed interviews of present and former chairpersons and a large sample of academic and service administrators, the study sought to present a practical look at the expectations and realities of the job of the academic department chairperson as it is, currently. Waltzer found that the job carries little formal authority, and that the authority that is posited in the job derives from what university and divisional administrators and department faculty allow in particular circumstances. Further, the chairpersons reported that increased bureaucratization and the prevailing styles of governance by councils, committees, and the like, spread authority to many other places and diminishes the authority available to the chairpersons.

A kind of condition or status accrues to the chairperson in which he or she has the responsibility for making the contradictory elements of effective,

efficient management and maintenance of collegial community function as a vital enterprise. Waltzer clarifies this condition by pointing out that the chairpersons must: (1) manage administrative directives and collegial decision making; (2) retain the friendship and respect of their colleagues while implementing policies that directly affect faculty; and, (3) accept responsibility for all departmental affairs but be one among equals in their departments (Waltzer, p. 14). This condition supports the need for performance assessment which is highly objective and free from political influences.

A question arises as to the value and importance of the chair position. In some colleges and universities the chairperson role is seen as one that is reluctantly held and/or one that may be rotated among senior members of a department. In many colleges and universities departmental responsibilities are shared in a highly collegial environment where the role and influence of the chair is minimal. In other organizations the chairperson is a powerful member of the organization with much potential influence.

It seems reasonable to assume that in the 1980's and beyond, chairpersons are likely to be expected by colleagues (including superiors) to be skilled in managerial functions, and, chairpersons are more likely to be evaluated according to managerial performance criteria than on the basis of purely academic performance criteria.

Millet (1978) regards the chairperson as a program planner and program manager. In the role of program planner the chairperson is concerned with providing department leadership in addressing such matters as student numbers, student quality, student advising, student performance and accomplishment of degree requirements. In the manager role the chairperson is working on tasks having to do with faculty personnel actions (recruitment, promotion, tenure,

separation), faculty budget actions (compensation, supplies, equipment, travel, departmental support), support personnel actions (recruitment, work assignment, etc.), faculty facilities (offices, classrooms, laboratories), and work scheduling (Millet, p. 53).

With the variety of functions, activities, and actors interacting with the role of chairperson, the need for multiple methods of assessment to verify performance is made prominent. It was generally anticipated that the kinds of task activities listed above would emerge from the assessment methods as the major domains or dimensions of performance. These activities certainly are not the exclusive listing of all such activities. Other tasks and functions such as representation of the department to external publics and one's personal professional performance as a scholar and teacher may be regarded as vital functions to be evaluated.

It appears that careful definition and a high degree of objectivity are characteristics which need to be part of the evaluation of academic department chairpersons. In a recent publication, Nordvall (1979, p. 14) points out that rating scales of performance related to characteristics such as those represented generally by leadership, interpersonal relationships, basic underlying traits, and commitment to institution seem to enjoy wide use. As typically developed and implemented (and sometimes borrowed from other institutions) the rating scale can be a device fraught with problems. Many evaluation rating forms tend to be disorganized and ambiguous. Evaluation rating scales may contain global behavior measures and vague trait descriptions. Rating scales are available for the evaluation of chairpersons, but according to Hodgkinson (1978, p. 110), "they are mostly opinion surveys and do not represent behavioral consensus on what excellent performance means."

Methods/Techniques

In order to overcome the shortcomings of rating scales and in order to establish a firm undergirding for the development and implementation of structured interviews and self-report guides, the methodology of the behaviorally anchored rating scale was chosen for use.

Based upon the work of Smith and Kendall (1963), and Harari and Zedeck (1973), Blood (1974), (1979) has demonstrated how a performance appraisal technique, behaviorally anchored rating scales (BARS), positively responds to many of the problems of evaluation identified above. The features of BARS are as follows:

1. A population of would-be raters are asked, individually, to write descriptions of behavioral episodes that define scale points. Hence, performance is defined in terms of observable behaviors by members of the population who later will do the rating. Borman and Dunnette (1975) address the point that the methodology has good potential for overcoming or reducing many of the errors often encountered in job performance rating systems. The involvement of superiors, subordinates, and/or job incumbents in all phases of the development of job behavior observation scales should enhance and facilitate the choosing of job dimensions and behavior examples that are readily understood and accepted by the persons asked to make the performance ratings. They indicate that by collecting critical incidents about job performance and then using them to define dimensions and to anchor different levels of performance on each dimension, the method should also help to decrease the

semantic ambiguities that tend to be prevalent in many performance rating systems. Because levels of performance are better defined one should anticipate decreased error attributed to leniency; because performance dimensions are better specified decreased halo effects should obtain; because raters are likely to be more attentive to the rating task it is likely that ratings assigned by different individuals will be congruent; and, because the scales help the raters focus directly on actual job behavior examples instead of traits it is likely that greater differentiation between persons being rated will result (Borman and Dunnette, p. 561).

2. The process (above) addresses salient performance dimensions. Faculty are used to construct BARS for evaluation of academic department chairpersons. A more detailed elaboration of the method could involve academic and support services administrators. Such an elaboration most likely would yield a measure of validation to overall performance appraisal as well as provide several new dimensions for assessment.
3. Meanings of response categories can be empirically verified. Blood (1979, p. 114) explains the double-elimination verification system as one where every behavioral episode generated is subjected to two judgments in our empirical sample drawn from the rater population (faculty). Each item, he says, is judged as to the performance dimension it represents, and all items of low agreement are dropped. The remaining pool of scale anchors then, consists of only those items which have

high commonality of meaning within the rater population.

4. All of the BARS generated are expressed in the language of the rater since the raters have actually generated the scales.

It appears that the methodology itself aids in clarification of goals, and by its participative and callaborative characteristics one could assume that the some of the basic values inherent in a collegial body are being taken into account.

Generation of Scales

The chairpersons of 14 academic departments in a college of arts and sciences were to be the target of the evaluation effort using the behaviorally anchored rating scales.

A total of 28 faculty members, two from each department, were asked to identify evaluative demensions for assessing chairperson performance. For each department, one of the faculty members selected to participate in the process had five or less years experience in the department, and the other faculty member had to have at least 10 years experience in the department. These criteria were achieved in each department. At a group meeting a brief definition was identified for each dimension after considerable discussion. This part of the process identified eleven performance dimensions (see Table 1). The faculty were then asked to provide critical/behavioral incidents that would demonstrate poor, adequate, and good performance on each of the eleven dimensions. Not all faculty provided three examples for each dimension. Instead of 924 (28 x 11 x 3) incidents, a total of 753 incidents were generated.

A three-member review panel examined the incidents and eliminated duplicate (redundant) incidents, non-behavioral episodes, and ambiguous episodes. This process resulted in a final count of 246 items. All of these items were

re-written into an "expectations" format. Each illustrative incident was stated in the form "could be expected to...," instead of remaining in a form which would imply that the chairperson to be rated actually had to exhibit the specific behavior. (Blood, 1979; Campbell, et. a.; Smith and Kendall).

A total of 42 (three from each department) different faculty was asked to make two judgments for the list of items. The first judgment required categorization of the 246 items into the eleven performance dimensions. The second judgment required placement of each item on a 7-point scale based upon the level of performance indicated. A rating of one represented the lowest level of performance and a rating of seven represented the highest level of performance. Any item for which agreement as to dimension represented was not at least 75 percent was eliminated from further consideration. This process of elimination forces a level of consensus on a final set of items.

The grouping of the final pool of items into dimensions yielded only seven of the original eleven dimensions since there were not enough items for constructing four scales (see Table 1). Blood (1979) says this circumstance can occur because: (1) some dimensions are not well-defined; (2) dimensions are similar to others; or (3) there is simply a low level of agreement as to the appropriateness of the behavior as specified.

Scales were then constructed for the remaining items. At least six items were used to anchor the meaning of scale points. The mean rating of the item located the item on the scale. By way of example, one of the scales is shown in Table 2.

The seven scales in their final form are to be distributed to all faculty of 14 academic departments of a collegeof arts and sciences in a state university. Members of this faculty nave participated in the generation of the rating

scales as indicated by the procedure thus outlined.

Results/Conclusions

In a recent publication on evaluation of administrative performance, Farmer (p. 18) points out that rating scales have several weaknesses although permitting ease of administration and anonymity. Weaknesses noted were biases introduced by: (1) friendship, (2) quick guessing, (3) appearance, (4) prejudices, (5) halo effects, (6) errors of central tendency, and (7) leniency. The behaviorally anchored scales proposed in this paper should respond positively in the elimination of most of the weaknesses identified. The method proposed is similar to that developed by Findlay College by Rasmussen (1978) although Rasmussen grouped scale items for dimension identification with factor analysis.

The scales, as applied, are to serve as a primary information source in a chairperson performance assessment system currently being implemented. Application and examination/analysis of behaviorally anchored rating scale results, interviews with faculty and key administrative staff, and self-reports of chairpersons, will enable the college of arts and science administrators to thoroughly assess the performance of the chairpersons and the behavioral specifications will not only afford the chairpersons meaningful feedback but may also indicate what kind of behavior should be demonstrated. Of course, a set of scales could be developed from the administrator point of view, as well. The information base generated by the activities identified here should be of much assistance in goal setting and in the definition of desired behavior.

REFERENCES

Blood, M.R. Spin-offs from behavioral expectation scale procedures. Journal of Applied Psychology, 1974, 59, 513-515.

Blood, M.R. Behavior-based teaching evaluation for specific educational programs. In the proceedings of a conference on The Assessment of Quality of Master's Programs, University of Maryland, College Park, 19791

Booth, D.B. Department and chairperson development. In New Directions for Higher Education, edited by Charles F. Fisher, pp. 71-82. San Francisco, Jossey-Bass, No. 22, 1978.

Borman, W.C. and Dunnette, M.D. Behavior based vs. trait oriented performance ratings: An empirical study. Journal of Applied Psychology, 1975, 60, 561-565.

Brown, J.D. Departmental and university leadership. In D.E. McHenry and associates, Academic departments. San Francisco: Jossey-Bass, 1977.

Campbell, J.P. Dunnette, M.D., Arvey, R.D., and Hellervik, L.V. The development and evaluation of behaviorally based rating scales. Journal of Applied Psychology, 1974, 59, 15-22.

Cohen, M.D. and March J.G. Leadership and ambiguity. Carnegie Commission on Higher Education, New York: McGraw-Hill Book Co., 1974.

Dressel, P.L. Handbook of academic evaluation. San Francisco: Jossey-Bass, 1976.

Farmer, C.H. Why evaluate administrators? In Administrator evaluation: concepts, methods, cases in higher education by C.H. Farmer, pp. 6-13, Richmond, Virginia: Higher Education Leadership and Management Society, 1979.

Green, J.L. and Stone J.C. Curriculum evaluation. New York: Springer Publishing Co., 1977.

Harari, O. and Zedeck, S. Development of behaviorally anchored scales for the evaluation of faculty teaching. Journal of Applied Psychology, 1973. 58, 261-265.

Hodgkinson, H.L. Administrators, evaluation, and the stream of time. In New Directions for Higher Education, edited by C.F. Fisher, pp. 107-113. San Francisco: Jossey-Bass, No. 22, 1978.

Millett, J.D. Professional development of administrators. In New Directions for Higher Education, edited by C.F. Fisher, pp. 51-58. San Francisco: Jossey-Bass, No. 22, 1978.

Nordvall, R.C. Evaluation and development of administrators (AAHE-ERIC Higher Education Research Report No. 6). Washington, D.C.: American Association for Higher Education, 1979.

Rasmussen, G.R. Evaluating the academic dean. In New Directions for Higher Education, edited by C.F. Fisher, San Francisco: Jossey-Bass, No. 22, 1978.

Smith, P.C. and Kendall, L.M. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 1963, 47, 149-155.

Waltzer, H. The job of academic department chairmen. Washington, D.C.: American Council on Education, 1975.

Table 1

Dimensions for Evaluating Chairpersons

*A. General Administration - management of department office, including record keeping and clerical staff.

*B. Resource Management - the manner and quality of allocating fiscal and other resources.

*C. Sensitivity to Faculty - the extent to which faculty needs are identified and addressed.

D. Planning - the demonstration of some better future and concept of what is desirable.

*E. Department Representation - the extent to which the chairperson interacts with publics internal and external to the university.

*F. Communication with Faculty - the degree to which faculty are kept informed of organization policies, regulations, plans, and procedures.

G. Department Organization - the extent to which the faculty is deployed and managed to attain department goals.

H. Interactions with Students - the extent to which quality students are recruited and advised; general supervision of graduate students.

I. Professional Development - the degree of encouragement and support given to individual faculty activity.

*J. Evaluation of Faculty - the extent to which various facets of faculty effort is evaluated.

*K. Curriculum Administration - the extent to which program features are monitored and modified.

---

*Dimensions identified by asterisk were retained in the final scales.

Table 2

General Administration

7	
	° you would expect this chairperson to have developed a complete set of office procedures and administrative forms
6	
	° you would expect this chairperson to maintain a set of statistics (information) about recruitment, attrition, grades, and placement of students
5	
	° you would expect this chairperson to have the clerical staff brought up-to-date on the processing of departmental requests for supplies, materials, etc.
4	
3	
	° you would expect this chairperson to rely on a personal system of information storage creating dependency on part of office staff
	you would expect this chairperson to be confused about the scheduling
2	° of work to be done in the office
	you would expect this chairperson to be unable to locate important
	° student records, or a faculty member's request for travel funds
1	